Communications of the Association for Information Systems

Volume 12 Article 46

December 2003

An Analysis of the Growth of Computer and Internet Security Breaches

Kallol Bagchi
University of Texas at El Paso, kbagchi@utep.edu

Godwin Udo
University of Texas at El Paso, gudo@utep.edu

Communications of the Association for Information Systems

AN ANALYSIS OF THE GROWTH OF COMPUTER AND INTERNET SECURITY BREACHES

Kallol Bagchi
Godwin Udo
Information and Decision Sciences Department
The University of Texas at El Paso
gudo@utep.edu

ABSTRACT

This study uses the modified Gompertz model and sparse data to analyze the growth rates of different types of computer and Internet-related crimes**.** The Gompertz model is an appropriate diffusion model because it is capable of modeling two opposite behaviors: (1) acts of attacks and imitation of attacks and (2) deterrence acts to prevent such attacks. In addition, this model can handle sparse data adequately. The model was used to analyze various types of attacks. The results indicated that growth patterns of computer and Internet crimes differ in growth patterns and that a relationship exists between occurrences of such security breaches and uses of certain security technologies. Thus, for example, financial fraud and denial of service are growing at a faster pace. The study also found, for example, that an increase in virus-related incidents does not necessarily increase anti-virus software use.

Keywords: Computer and Internet security breaches, Gompertz model, diffusion model, bad innovation, types of crimes, growth patterns of crimes

I. INTRODUCTION

Computer and Internet-related crimes show no signs of abatement. A 2003 survey conducted by the CSI/FBI reports that 75% of surveyed firms and agencies detected computer security breaches and acknowledged financial losses as a result of computer breaches [Power, 2003]. CERT/CC [2003] reports computer security vulnerabilities nearly doubled in 2002 with 2437 separate holes reported in 2001 and 4129 reported in 2002. Following the same trends, the number of reported incidents also increased significantly with 52,658 documented in 2001 and 82,694 in 2002. Through the continual monitoring of hundreds of Fortune 1000 companies, Riptech found that general Internet attack trends are showing a 64% annual rate of growth [http://www.riptech.com].

Neumann [1999] states that costs of cyber crime are difficult to measure; however, these costs are reasonably substantial and growing rapidly. Garg et al. [2003] attempted to quantify the financial impact of IT security breaches by using event-study methodology. They came to the same conclusion: IT breaches are extremely costly. Lukasik [2000] claims that cyber crime costs are essentially doubling each year. The problem becomes even more complicated when one

considers that these crimes are underreported. Ullman and Ferrera, [1998] mention that, according to FBI estimates, only 17 percent of computer crimes are reported to government authorities.

Previous studies that focused on computer or information systems security issues lack empirical results on how different these security breaches are from one another and what their growth patterns are. Such empirical studies are important because some attacks enormously and rapidly disrupt the Internet infrastructure for a length of time, thus resulting in millions of dollars in losses. For example, the “Code Red Worm” virus infected more than 250,000 systems around the globe in nine hours on July 19, 2001, and its estimated total global economic impact was as much as $2.6 billion [Householder et al., 2002].

The growth of computer and Internet security breaches can be studied from an innovation diffusion perspective [Rogers, 2003]. Innovation diffusion literature is usually concerned with good innovations and thus biased towards good innovations. The study of bad innovations such as security attacks can alert readers to the fact that innovations are not always good and what actions need to be taken to prevent such bad innovations. The present study uses the concept of bad innovations by using the modified Gompertz model [Pitcher et al., 1978] which is capable of capturing attack incidences as well as deterrent activities. Based on past experiences, it can be inferred that not all attacks deserve the same attention and not all attacks may show the same type of growth rate. It is important to know how these various crime rates are growing. This question needs to be investigated empirically. Although estimation with a sparse set of data at an earlier stage of growth is challenging, past studies proved it to be useful. In this paper, we focus on different types of attacks, how these evolved, whether different types of attacks evolved similarly, and how deterrence effects are working.

The study is preliminary in nature for a number of reasons. Literature is almost non-existent on this topic. Data on different types of security breaches are sparse [Power, 2002]. One of the most referenced studies of security breaches, the CSI/FBI computer crime and security survey by Richard Power, contains only a few years of recent data [1996-2002]. Modeling such security breaches during the early stages of data availability is difficult but extremely critical. Analysis with sparse data is, however, not uncommon in research literature. For example, marketing literature reports the forecasting of sales of new products with as few as five years of data [Mahajan and Peterson, 1985]. The dynamic behavior of hundreds of good innovations shows similar characteristics during the early phases of growth as observed across many types of products [Bass, 1969; Mahajan et al., 1985; Jepson, 1976]. Previous works on forecasting from early data with a small number of data points include Lawton and Lawton [1979], Tigert and Farivar [1981], Kalish and Lilien [1986], Wright, Uprichard and Lewis [1997]. Lilien et al. [1981] and Dalal et al. [1998] updated parameter estimates for a new product by using data on similar products or expert judgment in a Bayesian framework. Sultan et al.[1990] used meta-analysis-based prior information with a few data points on a new product to obtain more robust posterior estimates.

In the absence of prior information and data on Internet attacks, we use traditional diffusion models. Previous research reports that the shape of sales curves of many innovative products during the growth phase is similar [Mahajan et al., 1985]. Sales of new products in the early phases tend to grow extremely rapidly. This high growth rate tends to decrease over time and finally the diffusion matures and tapers off, as newer technologies replace older ones. Previous research also found that while exponential or logistic curves are adequate for modeling purposes in the growth phase, they are not adequate to model many innovations at an earlier stage. A small error at an early stage can result in a large effect on later time period forecasts [Martino, 1972].

Modified Gompertz curves, such as the General Sales Growth Curve [Lieb Associates, 2001], are reported which describe the data well and yield good curve fitting and forecasting of new innovations in the early growth phases [Jepson, 1976; Lakhani, 1979]. The Gompertz curve could be a good fit for innovations which rapidly increase in the beginning and then taper off slowly. The point of inflection of the growth curve occurs at 33% of total potential diffusion. Such a model is

used in the present study of bad innovations [Pitcher et al., 1978]. In addition, the model's explanatory power helps to understand how these attacks are developing and what factors are behind such attacks.

II. TYPES OF BREACHES

Some of the important security breaches since 2001 are the results of the following attacks [CERT/CC, 2003]:

- Multiple vulnerabilities in the Internet Software Consortium's Berkeley Internet Name Domain (BIND) server,
- Sadmind/IIS worm (a worm that exploits a vulnerability in Solaris systems and subsequently installs software to attack Microsoft IIS web servers),
- Code Red worm (a self-propagating malicious code that exploits IIS-enabled systems),
- SirCam worm (a malicious code that spreads through email and potentially through unprotected network shares), and
- Nimda blended threat (a combination of worm, viruses, and other codes that propagates itself via several methods, including email, network shares, or through an infected web site).

Security breaching techniques have come a long way from early hacker-induced attacks of 1970s. Sophisticated attacks include superior software techniques that are increasingly difficult to separate from normal network traffic. An automated sophisticated attack may typically consist of four phases:

1. scanning for potential victims,
2. compromising vulnerable systems,
3. propagating the attack, and
4. coordinated management of attack tools [Householder et al., 2002].

To increase attack efficiency, scanning and attack tools are integrated and attack cycles are initiated automatically. Distributed attack tools are common.

The main types of reported popular Internet-based attacks are [Denning, 1999; Smith and Rupp, 2002; Ratnasingam, 2002; McCroham, 2003]:

• denial of service	• domain name system	• web defacement
• worm and virus	• router attacks	

The denial of service attack prevents legitimate users from using the service, typically by flooding a network or disrupting connections or services. An example is the Mafia Boy attack from February 7-9, 2000 on web sites such as Yahoo.com, CNN.com, and Amazon.com. The web sites went out of service for more than two hours costing $1.2 billion in loss in business [CCITS, 2002].

A worm is a self-propagating malicious piece of code and is highly automated. A recent example of a worm is the “Blaster” or “Sobig” worm which affected over 500,000 computers in September 2003 [Krebs, 2003]. Top-level domain servers are potentially vulnerable and any attack on them can cause widespread problems.

Some viruses can be spread by executing infected programs. An example of a destructive virus is “I love you,” which appeared in May 2000, took five hours to spread, and cost some $10 billion in damages and lost productivity [CCITS, 2002]. Sometimes, when an infected program runs, it may wipe out the hard disk and do other damages.

Routers, which are devices used to direct traffic on a network, can be attacked in several ways such as by denial of service or by using the router as an attack platform. It is no surprise that many corporate network professionals cite e-mail parasites (62%) and spam (17%) as the two most damaging types of external security attacks.

The CSI/FBI report [2002] talks about many other types of computer and Internet-based attacks or misuses such as financial and telecom fraud, telecom eavesdropping, sabotage, laptop misuse, active wiretap, and insider abuse of net access. Web defacement is treated as a separate type of attack because of its importance and recent frequency (more than 50 a day in 2002, [CTNEWS, 2002]. Reasons for web defacement include electronic graffiti, attention seeking, and intellectual challenge. Domains such as .gov, .mil, .com are frequently targeted for web defacement attacks. Mirror web sites such as Alldas.de, attrition.org, and safemode.org chronicled this phenomenon and were also closed down by hacker attacks.

III. SECURITY BREACH DETERRENT TOOLS AND TECHNIQUES

As security breaching techniques refine, so also do security attack detection and prevention techniques. Several tools are available to firms to combat security attacks. Firewalls, placed between the company network and the Internet, provide ongoing protection by denying suspicious traffic. Another system, called the intrusion detection system (IDS), is needed to inform companies when they are under attack. The IDS examines all packets and prepares a log file. The security administrator examines the log file to look for suspicious patterns and generates messages for possible attacks. If an attack packet passes through the firewall, the next line of defense is to prepare the host from possible attacks by installing vendor-specific current patches for known weaknesses in the system. A large number of attacks emerge from known weaknesses in popular software. Security systems are also designed to prevent eavesdropping attacks. Secure communication is ensured when the checks for authenticity, integrity and confidentiality are maintained. Many techniques such as biometrics, digital IDs, encrypted logins, anti-virus software, and access control mechanism are used to prevent attacks [Power, 2002]. Not all of them are universally effective or popular.

Sometimes, an attacker succeeds by breaking all systems. This situation is called a security incident. Companies need to plan for incident handling (also called incident response). A good plan will detail how to stop the attacks, restore the system to its pre-attack state, and how to document that attack for future prosecution. In case a firm’s security administration fails, Internet security sites can provide help. Organizations such as CSI, CERT, NIPC, and the IEEE task force on security and privacy [IEEE, 2002], make enormous deterrence efforts to stop hacking that maliciously damages academic, government, and business activities.

The security infrastructure and security providers no doubt act as a deterrent to attempts of such breaches by sustained organized efforts. Security laws and regulations of a nation also help the deterrent side of the equation. Some U.S. government regulations are already in place. These regulations are related to computers, access devices and communication lines, stations, and systems. For example, the computer fraud and abuse statute 1030 states that if anyone knowingly or intentionally accesses a computer without authorization or exceeds authorized access, he/she is liable to be punished [NSI, 2002]. International efforts are also not lacking. Forty-one European countries, plus the U.S., Canada and Japan, attended a recent convention

on cyber crime. These nations signed a treaty that supplies a legal framework aimed at protecting society against cyber crime [Conventions, 2002].

IV. A MODEL OF SECURITY BREACHES/ATTACKS

Many researchers have studied Computer/IS security issues [Atkins, 1996; Parker, 1983; Straub, 1990]. Straub and his coworkers used general deterrence theory in the IS environment [Straub, Carlson, and Jones, 1993; Straub, 1990; Straub and Welke, 1998]. The basic argument in this body of work is that information security actions can deter potential computer abusers from committing illegal acts. They also found empirical evidence that security actions can lower systems risk.

Arquilla [2001] argues that the information revolution favors the rise of network forms of organization to redefine societies, and in so doing invites the duel between conflict and cooperation. The term 'netwar' calls attention to the prospect that network-based conflict and crime will be major phenomena in the years ahead. Arquilla thinks that the spread of the network form and its technologies is clearly bringing some new risks and dangers. It can be used to generate threats to freedom and privacy. New methods for surveillance, monitoring, and tracking are being developed.

Mostly though, previous studies lack empirical results on how different types of attacks grow or provide reliable models of such attack growths. This understanding is important. Some attacks enormously and rapidly disrupt the Internet infrastructure for a length of time, thus resulting in millions of lost dollars. For example, the infamous "Melissa" virus in 1999 infected thousands of computers with rapid speed, causing an estimated $80 million in damages [CCITS, 2002].

The growth process can be studied from an innovation diffusion perspective [Rogers, 1991]. The four main elements in the diffusion process are:

1. the innovation (good or bad),
2. channels of communication,
3. time, and
4. the social system.

In the present case, examples of channels could be direct word-of-mouth or contacts made via the Internet/Web. Timer relate to the rate at which the innovation is diffused, and the social system is the system of all potential and existing attackers.

Ideally, a growth model is needed that can capture both deterrence and imitation activities to model the security breaching incidents. However, traditional diffusion models do not provide the necessary explanatory power to analyze the attack phenomenon adequately [Mahajan and Peterson, 1985].

V. THE GOMPERTZ MODEL

The modified Gompertz model used by Pitcher et al. [978] assumes that the probable causes for the outbreak of such incidents are imitative as well as inhibitive in nature. This model's theoretical background is strong and is based on a social conflict theory. The imitative aspect is based on incident news spread via the Internet and by word-of-mouth; the inhibitive aspects can also be spread via Internet/Web sites and related stories. However, people only engage in security attacks when they feel threatened or are motivated by some economic or other gain and observed the success of earlier attackers [Bandura, 1986]. Traditionally, the challenge or threat to such attackers was mostly an intellectual one: to break a system. To quote a hacker expert,

> *"It's the sheer challenge (to crack a code or break a system) rather than any (criminal intent). They see it as an intellectual challenge and a prize, (and) they look at the success of what they have done rather than the consequences of the lives of people they have affected"* [Dreyfus, 2002].

Of course, other types of challenges come, for instance, from making money or taking economic or political advantage. The more successful the earlier attackers are, the more aggressive the behavior of the present attacker becomes. Each such incident is an imitation of previous behavior and a behavioral model for others to imitate. On the other hand, the increase of security activities and success stories about preventing such attacks could reduce the number of attacks. Thus, a combination of imitation and inhibition as assumed by the asymmetric model could provide a realistic background in modeling such incidents.

The model can be expressed as:

$$\frac{dN(t)}{dt} = c \cdot e^{-qt} \cdot N(t)$$

where t = time,

N (t) = cumulative number of attack incidents at time t

c, q are parameters of the models.

The parameter c denotes the net rate of instigation to attacks and q denotes the rate of inhibition in such attacks.

We model the growth process as a combination of attack influences as well as preventive efforts by various agencies to curb such incidents. Our analysis suggests that the growth was indeed influenced by a combination of factors: attacks by like-minded peers (hackers or crackers) and attack-preventive measures put forward by various governments, academic, and security agencies. The implications of the results affect everyone - from security professionals and merchants associated with on-line trade over the Internet to academics, professionals, and other day-to-day users of the Internet/Web.

VI. PROPOSITION FORMULATION

Although imitative and deterrence acts constitute the background of any attack scenario, the rates of imitation and deterrence may not be the same. When the rate of instigation increases it may mean an overall increase in deterrence rate as more and more security products will be developed. As these products come onto the market, attackers find ways to bypass these products and refine their attacks, which in turn leads to more refined security products. This cycle of reinforcing attack and deterrence continues [Pitcher et al., 1978].

Proposition 1: Relative increase in net instigation rate is related to relative increase in deterrence rate.

Sofaer and Goodman [2001], observe that

> *"the risks of cyber terrorism and cyber crime vastly outweigh our abilities to control those risks by technological means, although technology can help and should be vigorously pursued."*

Thus, preventive measures are assumed to be thoroughly outweighed by attacks. Therefore, it is expected that the value of c, the net rate of instigation will be much higher than the value of q, the rate of deterrence or inhibition.

Proposition 2: Values of the net rate of instigation, c, will be much higher than values of q, the rate of inhibition for computer and Internet-related bad innovations, i.e., digital crimes and security breaches.

Although reported computer crimes are of many types, not all of them are equally popular, due to economic, political, technical and a variety of other reasons. At the beginning, hacking was done primarily for intellectual satisfaction, to break a system. In recent times however, financial profit considerations are one of the main reasons for computer crimes. Thus, losses in 2002 because of financial fraud and theft of proprietary information far surpass any other type of loss [Power, 2002]. Virus and insider net abuse still continue to cause concern and denial of service and system penetration incidents are rising rapidly. Laptop theft incidents decreased in recent years. It is expected, therefore, that the growth rates of crime technologies will differ by crime types.

Proposition 3: Not all computer crimes and security breaches show similar growth rates.

Security tools or defensive cyber weapons include encryption, authentication, access controls, firewalls, anti-viral software, audit tools, and intrusion detection systems [Denning, 2000]. Although new security tools are being developed (for example, biometrics and digital IDs) and security technologies are increasingly used by many firms, it is useful to investigate whether and how usage is related to attacks that occur. Thus, denial of service attacks, proprietary information theft, and system penetration attacks should lead to more use of intruder detection software, encryption, and firewalls; virus attacks should lead to more antivirus software use and encryption.

Proposition 4: The more security incidents happen, the more security technologies are used.

In particular,

> 4a. system penetration attacks should lead to more use of intruder detection software, encryption, and firewalls
>
> 4b. denial of service attacks should lead to more use of intruder detection software, encryption, and firewall
>
> 4c. proprietary information theft should lead to more use of intruder detection software, encryption, and firewalls
>
> 4d. virus attacks should lead to more antivirus software use and encryption use.

VII. DATA ANALYSIS AND METHOD

Power [2002] gathered data on aspects of cyber crime from 1997 onwards. The survey, called the "Annual CSI/FBI Computer Crime and Security Survey," gathers data based on a survey questionnaire sent to information security practitioners in U.S. organizations. The statistical rigor of the survey findings is sound [Richardson, 2003]. The same survey was used for all years and the respondents are mostly security professionals with firms who are better-informed about the attacks than others. This data set is used for the present study. We used total annual loss data which was available in U.S. dollar value. We decided to use the annual financial loss data because it would reflect the most accurate situation in terms of extent of financial damages incurred by firms by such attacks. The financial loss data were converted to 1996 U.S. dollar value by dividing by the price deflator for each year.[1]

For the 2002 survey, questionnaires were distributed to 3,500 information security professionals with a 14% response. The responses were anonymous. Job titles of respondents ranged from

[1] Alternatively, the 1996 numbers (first quarter) can be multiplied by 1.126 to show values in 2003 (first quarter) or multiplied by 1.108 to show values in 2002 (first quarter).

corporate information security manager and data security officer to senior systems analyst. Organizations surveyed included corporations, financial institutions, government agencies and universities. Total dollar losses as listed in several years are shown in Table 1. In terms of dollar value, theft of proprietary information and financial fraud caused the greatest financial losses.

Table 1. Total Dollar Losses (in millions) per Year
From Digital Crimes as Reported by Respondents

Year	No. of Respondents with Losses	Amount in US $ Value
1997	249	100,119,555
1998	241	136,822,000
1999	163	123,779,000
2000	273	265,337,990
2001	196	377,828,700
2002	n/a	455,848,000
2003	251	201,797,340

Source: [Richardson 2003]

Table 2 shows the distribution of the respondents who reported attacks by industry sectors for 2000 and 2003. The distribution did not change much over the four year period.

Table 2. Distribution of Respondents Reporting Attacks
by Industry Sectors

Industry Sector	%(Year 2000)	% (Year 2003)	Industry Sector	%(Year 2000)	% (Year 2003)
State Govt.	7	5	Transportation	2	1
Local Govt.	2	3	Telecom	4	4
Federal Govt.	9	7	Financial	17	15
Education	5	5	Manufacturing	10	11
Retail	4	3	Utility	4	4
Medical	7	8	Legal	--	1
High-tech	17	17	Other	12	17
			Total	100	100

Source: [Richardson 2003]

For web-defacement, we use monthly defacement incident data available from January 1997 to July 2001, based on a report by Attrition, a mirroring firm, [Attrition, 2001]. Attrition began actively mirroring defaced sites since 1995. However, it had to close down in 2002, due to hacker attacks. Sample data from web defacement are shown in Table 3.

Table 3. Sample of Monthly Web Defacement Incidents from 1995-2001

Month/Year	No of incidents
November 1995	1
November 1996	6
November 1997	2
November 1998	35
November 1999	665
November 2000	722
May 2001	1137
Grand Total (1995-2001)	15203

Source: www.attrition.org

The data analysis method used in this paper is a non-linear least square regression scheme. We used SPSS to design and run the non-linear model described above, with different sets of data. Non-linear equations are sometimes known to be difficult to converge. The convergence problem is handled with suitable initial values of parameters. The Levenberg-Marquardt algorithm [SPSS, 2003] is mostly used to determine parameter values of interest, q and c here. In a few cases, we needed to constrain the parameter values to obtain a solution.

VIII. RESULTS

Proposition1. Relative increase in net instigation rate is related to relative increase in inhibition rate. Figure 1 shows the result for Proposition 1. The break-even line in the figure shows that there is no point above the line, thus showing that net instigation rates are always higher than the inhibition rates. The figure captures the fit of the power function of the relationship between q and c. The function is: $q = .089c^{(2.19)}$ ($R^2 = .66$). An increase in net instigation rate is greater than the corresponding relative increase in inhibition rate. This result is consistent with results obtained from other types of crimes [Pitcher et al., 1978]. The moderate fit and the positive value of c support Proposition 1.

Proposition 2. Values of net instigation rate, c, will be much higher than values of inhibition rate, q, for computer and Internet-related bad innovations, i.e., computer crimes and security breaches. Table 4 presents results from running the model for various types of computer crimes and security breaches. The R^2 value from the model fits are high (.80-.99). The values of q and c are different, for each type of security breach, with values of c much higher than q. When c> q, overall impact of net instigation is more than the inhibition rate and vice versa. The results are again consistent with the results obtained from other types of crimes [Pitcher et al., 1978]. Proposition 2 is confirmed.

Proposition 3. Not all computer crimes and security breaches show similar growth rates. The results are shown in Table 4. The pair of values of q and c, as obtained from each run, is very different for each type of crimes, thus confirming Proposition 3. Of these viruses, financial fraud, and theft of proprietary information are projected to be significant and costly in the near future. Denial of service is rising rapidly. Telecom fraud, active wiretapping, laptop theft, and unauthorized insider access will be lower. By comparison, the rest of the crimes are projected to be at a moderate level of intensity.

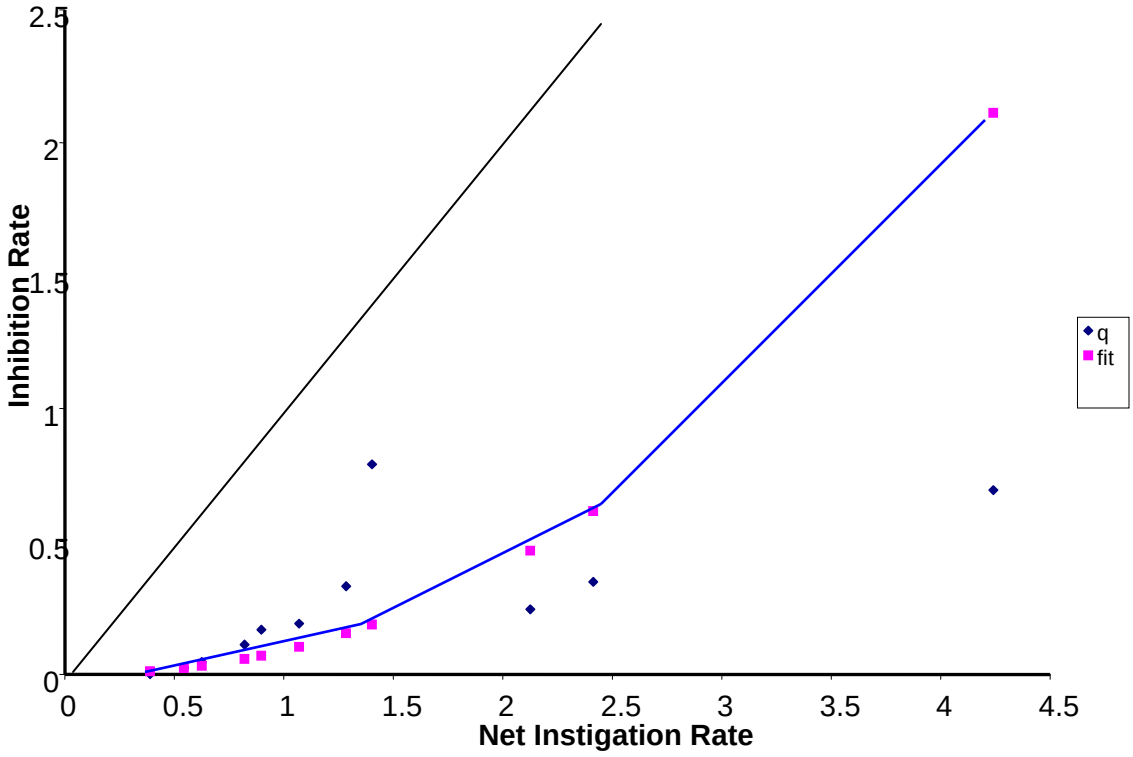

Figure 1. Inhibition Rate (q) vs. the Net Instigation Rate (c)

Table 4. q and c Values from the Model of Various Attack Types

Items	Upper limit cumulative cost	R^2	q	c
Theft of proprietary info.	5.9E+09	0.99	0.1555	1.029
Sabotage of data of networks	2.6E+09	0.95	0.1117	.8208
Insider abuse of Net access	9E+08	0.99	0.245	2.13
Financial fraud	1.02E+13	0.99	0.046	.626
Denial of service	2.79E+08	0.99	0.191	1.07
Virus	3.42E+16	0.98	0.024	.544
Unauthorized insider access	7.75E+07	0.89	0.693	4.24
Telecom fraud	5E+07	0.94	0.79	1.403
Active wiretapping	2.3E+07	0.80	0.348	2.41
Laptop theft	8.8E+07	0.99	0.331	1.255

[1]unit is number of incidents

Proposition 4. The more attack incidents happen, the more security technologies are used. The results are shown in Table 5. Limited support is found for propositions 4a-4c; whereas proposition 4d is falsified.

As Table 5 indicates, more denial of service and system penetration attacks increase use of encryption and intrusion detection software. More financial frauds and theft of proprietary information lead to increase in anti-virus software and firewall use. However, unexpectedly, an increase in virus-related incidents does not increase anti-virus software or other security software use. Power [2002] reported that although 90% of respondents used anti-virus software, 85% of them were affected by virus, worm, and other attackers. Inadequacy of existing software in combating viruses is evident to managers, as viruses always come in newer forms. More research is needed on this issue.

Table 5. Pearson Correlation Coefficients and Standard Deviations for Security Tools/Techniques and Attack Types

		Security Technologies Used by US Firms (1998-2002)			
Types of Attacks or Misuse Detected	Measure	% Antivirus Software Used	% Intrusion Detection Software Used	% Encrypted Login Used	% Firewalls Used
Denial of Service	Pearson Correlation	.145	.973	.953	.635
	Sig. (2-tailed)	.816 (N.S)	**.005****	**.012***	.250 (N.S)
Financial Fraud	Pearson Correlation	.854	-.450	-.287	.482
	Sig. (2-tailed)	**.065**	.447 (N.S)	.639 (N.S)	.411 (N.S)
System Penetration	Pearson Correlation	.204	.985	.953	.675
	Sig. (2-tailed)	.743 (N.S)	**.002****	**.012***	.211 (N.S)
Theft of Proprietary Info	Pearson Correlation	.833	.381	.609	.884
	Sig. (2-tailed)	**.08*****	.526 (N.S)	.275 (N.S)	**.047***
Virus	Pearson Correlation	.365	.481	.671	.534
	Sig. (2-tailed)	.546 (N.S)	.412 (N.S)	.215 (N.S)	.354 (N.S)

N.S—Not significant; * Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).; *** Correlation is significant at the 0.1 level (2-tailed)

IX. DISCUSSION

How robust is the model fit using the sparse data model (Gompertz in the present study)? We could check only for Web defacement cases, as it had enough monthly data for forecasting purpose. Monthly data are more susceptible to fluctuations than yearly data. We used 12 data points for building the model (starting from July 1998, as early data were not contiguous) and 12 successive data points for prediction, using the Gompertz model. Figure 2 shows the results.

As shown in Table 6, for the first 12 forecasts, predicted value exceeded (in three cases) 50% of the actual value. Thus, 30% of forecasts were off by more than 50% (refer to the last column of the table). The average error for the 12-month forecast was 32% and the maximum error in forecast during this time frame was 74%. Predictions up to 11 more months could be observed (i.e., up to the end of available data) in which predictions exceeding 75% greater than actual values (but less than100%) were 10 in number, or roughly 44% of total number of predictions.

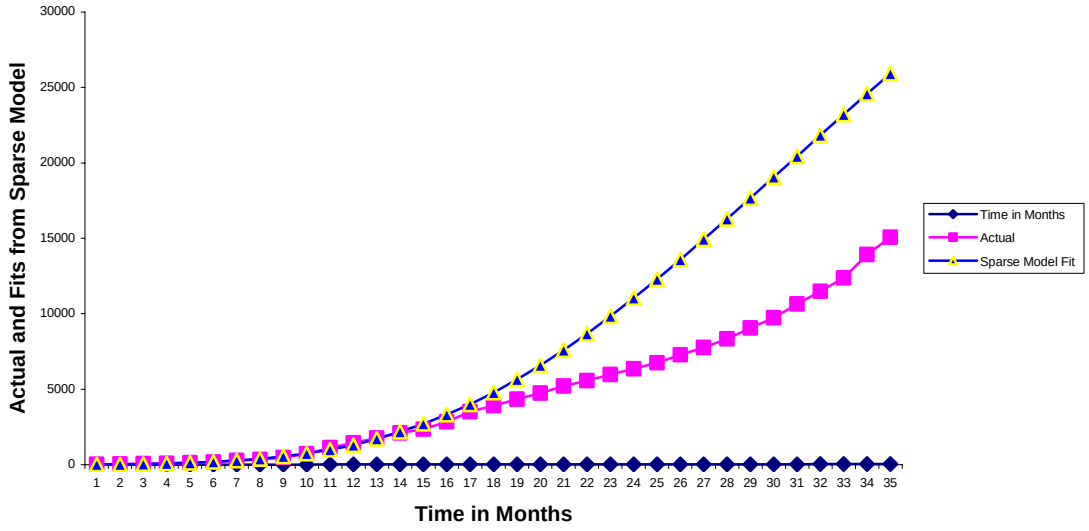

Figure 2. The Web Defacement Data and Predictions from the Gompertz Model

Table 6. Sample Web Defacement Data and Sparse Model-Based Forecasts

Forecast No.	Actual Incidents (Cumulative)	Gompertz Forecast	Forecast/ Actual	Forecast No.	Actual Incidents (Cumulative)	Gompertz Forecast	Forecast/ Actual
1	1762	1716	0.97	13	6746	12293	1.82
2	2088	2171	1.04	14	7288	13592	1.87
3	2346	2703	1.15	15	7764	14926	1.92
4	2843	3315	1.17	16	8337	16286	1.95
5	3508	4009	1.14	17	9059	17665	1.95
6	3920	4786	1.22	18	9742	19054	1.96
7	4345	5645	1.30	19	10639	20446	1.92
8	4740	6583	1.39	20	11487	21835	1.90
9	5203	7596	1.46	21	12374	23213	1.88
10	5559	8680	1.56	22	13920	24575	1.77
11	5968	9829	1.65	23	15057	25915	1.72
12	6347	11036	1.74				
		Average Error (%)	*32%*				

Note: Forecast 1 is for July 1999; Forecast 23 is for May 2001

These results confirm that the Gompertz model is adequate for short term predictions of web defacement incidents.

The exponential or the logistic models performed much worse, when compared with the Gomperz model, as shown in Table 7. Figure 3 graphically shows how the logistic model overestimates the data compared to the Gompertz model in the present case. Although the logistic model performed better than the exponential model, neither model was a good fit [2].

[2] We assumed the upper bound of the logistic model as 100,000 which is a little more than six times the last cumulative actual value (as obtained in May 2001). This is a conservative, low assumption as the actual

Table 7. Sample Web Defacement Data and Fits

Time (in Months)	Actual Cumulative Incidents	Sparse Model/ Gompertz	Exponential Model	Logistic Model
(July 1998) 1	12	17	22	22
2	37	29	32	32
3	63	48	48	48
4	83	78	71	71
5	118	121	105	106
...	..	..	..	..
30	9742	19054	1958410	95283
31	10639	20446	2901758	96771
32	11487	21835	4299509	97800
33	12374	23213	6370544	98506
34	13920	24575	9439177	98988
(May 2001) 35	15057	25915	13985944	99316

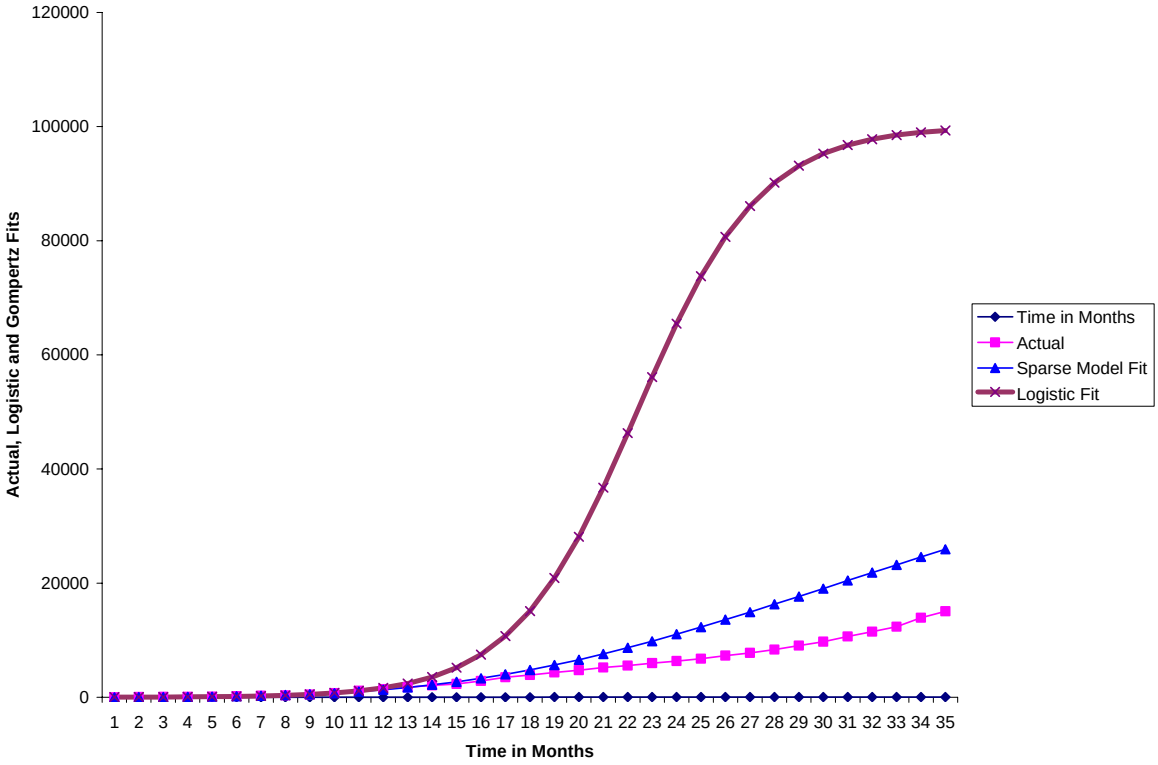

Figure 3. A Comparison of Logistic and Gompertz Fits For the Web Defacement Data

To verify this phenomenon for other types of attacks, we calculated absolute sum of errors based on the fits obtained from logistic, exponential and Gompertz models for all attack types. The

growth registered has been more than 1200 times the initial starting value in July 1998 (i.e., see Table 7). Higher values of upper bound may lead to higher errors in estimates and lower values may be impractical.

results are shown in Table 8. The Gompertz model provides superior fits for this class of models in all but two cases.

Table 8. The Fits from Competing Models

Rank of Absolute Sum of Errors				
Items	Gompertz	Logistic	Exponential	Model That Provides the Best Fit
Theft of proprietary info.	2	1	3	Logistic
Sabotage of data of networks	1	2	3	Gompertz
Telecom eavesdropping	1	3	2	Gompertz
Insider abuse of Net access	1	3	2	Gompertz
Financial fraud	1	2	3	Gompertz
Denial of service	1	2	3	Gompertz
Virus	1	2	3	Gompertz
Unauthorized insider access	1	3	2	Gompertz
Telecom fraud	1	2	3	Gompertz
Active wiretapping	3	1	2	Logistic
Laptop theft	1	2	3	Gompertz
Web defacement	1	2	3	Gompetrz

Note. Absolute sum of errors is ranked 1, 2 or 3 based on minimum values

X. CONCLUSIONS

Of the four propositions explored in this study, three (Propositions 1-3) were strongly confirmed while the remaining one (Proposition 4) was partially confirmed. In summary, the results of this study led us to conclude that:

- Relative increase in net instigation rate is related to relative increase in inhibition rate which implies that the increasing attack incidences will force organizations and governments to come up with means of preventing or reducing them;
- For computer and Internet-related attacks (bad innovations), the values of net instigation rate is higher than values of inhibition rate, implying more efforts and resources need to be applied toward inhibiting attacks;
- Different computer crimes and security breaches grow at different rates, which implies that all these crimes should not receive the same level of attention because some crimes are likely to spread more rapidly than others;
- Real world practice does not always follow the common notion that as more attack incidents occur, more security technologies are used. This finding may imply that organizations and governments do not necessarily spend money on security measures in proportion to the frequency of attack incidences. Ninety percent of respondents in the 2002 survey, for example, used anti-virus software; however, at least 10-15% of respondents did not detect any virus, due probably to non-use or ignorance [Power, 2002]. Viruses are among those attack incidents that caused financial losses.

This article is a first attempt to identify the nature of growth of various computer and Internet-related crimes, using a sparse set of data. First, a model was selected for bad innovation modeling which can represent both imitative and inhibitive behaviors in attacks. Next, the model was used to derive and compare various types of attack statistics with a sparse set of data.

Although the modified Gompertz model used is better than other forecasting models, it may still yield forecast errors that can only be refined with the progress of time.

Future predictions can be richer if the underlying relationships of the regression model remain unchanged. We did not use the data from 2003 as it trended downwards compared to the trend of the overall set (1997-2002). Nonetheless, the results should be interpreted with caution. However, our objective is to obtain and compare preliminary growth estimates of various attacks and this paper indicates the different rates at which such crimes are growing.

Editor's Note: This article was received on October 8, 2003 and was published on December 19, 2003. It was with the authors for three weeks for two revisions.

REFERENCES

EDITOR'S NOTE: The following reference list contains the address of World Wide Web pages. Readers who have the ability to access the Web directly from their computer or are reading the paper on the Web, can gain direct access to these references. Readers are warned, however, that

> 1. these links existed as of the date of publication, but are not guaranteed to be working thereafter.
>
> 2. the contents of Web pages may change over time. Where version information is provided in the References, different versions may not contain the information or the conclusions referenced.
>
> 3. the authors of the Web pages, not CAIS, are responsible for the accuracy of their content.
>
> 4. the authors of this article, not CAIS, are responsible for the accuracy of the URL and version information.

Arquilla, J. (2001) *Networks and Netwars: The Future of Terror, Crime, and Militancy,* Santa Monica, CA: RAND Corporation.

Atkins, D. (1996) *Internet Security Professional Reference*. Indianapolis, IN: New Riders Pub.

Attrition (2001) http://www.attrition.org/mirror/attrition/stats.html Last consulted 11-27-03.

Bass, F.M. (1969) "A New Product Growth Model for Consumer Durables" *Management Science* (15), pp. 215-227.

Bandura, A. (1986) *Social Foundations of Thought and Action*. Englewood Cliffs, NJ: Prentice-Hall.

A CCITS/Infosech Presentation on Internet Security, 2002.

CERT/CC Web Site, http://www.cert.org . Last consulted 12-06-03.

Convention on Cybercrime http://conventions.coe.int/Treaty/EN/cadreNews.htm . Last consulted 12-06-03.

Computer Fraud and Abuse Statute, (2002) http://nsi.org/Library/Compsec/cfa.txt . Last consulted 12-06-03.

CTNEWS(2002) http://www.cnetnews.com. Last consulted 11-06-03.

Dalal, S., Ho, Y. and Sherman, R. (1998) "Learning from Experience to Improve Early Forecasts: A Posterior Mode Approach" In Business and Economic for the 21st Century, Vol. II Worcester, MA: Business and Economics Society International,. pp. 338-353.

Denning, D. (2000) "Reflections on Cyberweapons Controls" *Computer Security Journal,* (XVI) 4.

Denning, D. (1998) *Information Warfare and Security*, Upper Saddle River, NJ: Pearson Education.

Dreyfus, S. (2002) “Cracking the Hackers' Code” http://www.smh.com.au/articles/2002/08/20/1029114072039.html. Last consulted 12-06-03.

Ford, R. (1999) “No Surprises in Melissa Land” *Computers and Security*, (18), pp. 300-302.

Garg, A., Curtis, J. and Halper, H. (2003) “Quantifying the Financial Impact of IT Security Breaches” *Information Management & Security* (11)2, pp. 74-83.

Householder, A., Houle, K. and Dougherty, C. (2002) “Computer Attack Trends Challenge Internet Security, Security and Privacy” Supplement to Computer, *IEEE Computer Society*.

Jepson, C., *E. I. DuPont de Nemours & Co., Inc,* Internal Presentation, 1976.

Kalish, S. and Lilien, G. (1986) “A Market Entry Timing Model for New Technologies” *Management Science*, *32* (2), pp. 194-205.

Katz, M. and Shapiro, C. (1986) “Technology Adoption in the Presence of Network Externalities” *Journal of Political Economy* (94), pp. 822-841.

Krebs, B. (2003). “Good' Worm Fixes Infected Computers” http://www.washingtonpost.com/wp-dyn/articles/A9531-2003Aug18.html. Last consulted 12-06-03.

Lakhani, H. (1979) "Empirical Implications of Mathematical Functions Used to Analyze Market Penetration of New Products" *Technological Forecasting and Social Change* (15)2, pp. 147-156.

Lawton, S. B. and Lawton, W. H. (1979) “An Autocatalytic Model for the Diffusion of Educational Innovations” *Educational Administrative Quarterly*, *15* (1), pp. 19-46.

Lukasik, S. J.(2000) “Protecting the Global Information Commons" *Telecommu-nication Policy*, (24)6-7, pp. 519-531.

Mahajan, V., Muller, E. and Bass, F. M. (1990) "New Product Diffusion Models in Marketing: A Review and Directions for Research" *Journal of Marketing*, (54), pp. 1-26.

Mahajan, V. and Peterson, R. (1987) “Models for Innovation Diffusion,” Sage University Paper series on Quantitative Applications in the Social Sciences, (2nd Ed.), Beverly Hills: SAGE Publications.

Martino, J. P. (1972) "The Effect of Errors in Estimating the Upper Limit of a Growth Curve" *Technological Forecasting and Social Change*, (4), pp. 77-84.

McCrohan, K. F. (2003) “Facing the Threats to Electronic Commerce” *Journal of Business & Industrial Marketing***,** 18 (2) , pp. 133-145.

Neumann, P. (1999) “Information System Adversities and Risks” presented at the Conference on International Cooperation to Combat Cyber Crime and Terrorism, Stanford, CA: Hoover Institution, , pp. 1-2, 3. http://www.oas.org/juridico/english/information_system_adversities_a.htm

Parker, D.B. (1983) *Fighting Computer Crime*. New York: Scribner's.

Pitcher, B., Hamblin, R. and Miller, J. (1978) “The Diffusion of Collective Violence” *American Sociological Review*, (43), pp.23-35.

Power, R. (2002) “CSI/FBI Computer Crime and Security Survey” *Computer Security Issues and Trends*, (8)1, pp. 1-22.

Ratnasingam, P. (2002) “The Important of Technology Trust in Web Services Security” *Information Management & Computer Security*, (10)5, pp. 255-260.

Richardson, R. (2003) The 2003 CSI/FBI Computer Crime and Security Survey. San Francisco**:** Computer Security Institute Inc., pp. 1-20.

Rogers, E. (2003) *The Diffusion of Innovation*. New York: Free Press.

Smith, A. D. and Rupp, W. T. (2002) "Issues in Cybersecurity: Understanding the Potential Risks Associated with Hackers/Crackers" *Information Management & Computer Security*, (10)4, pp. 178-183.

Sofaer, A. D. and Goodman, S. (Eds), (2001) *The Transnational Dimension of Cyber Crime and Terrorism,* Hoover National Security Forum Series, Stanford, CA: Hoover Institution Press.

SPSS 11 Syntax Reference Guide, 2003. Chicago IL: SPSS Publication Sales.

Straub, D.W. (1990) "Effective IS Security: An Empirical Study'' *Information Systems Research*, (1)3, pp. 255-276.

Straub, D., Carlson, P. and Jones, E. (1993) "Deterring Highly Motivated Computer Abusers: A Field Experiment in Computer Security" *Journal of Management Systems* (5)1, pp. 33-48.

Straub, D. and Welke , R. (1998) "Coping with Systems Risk: Security Planning Models for Management Decision-Making" *MIS Quarterly*, (22)4, pp. 441-469.

Sultan, F., Farley, J., and Lehmann, D. (1990) "A Meta-Analysis of Applications of Diffusion Models" *Journal of Marketing Research*, (27), pp. 70-77.

Tigert, D. and Farivar, B. (1981) "The Bass New Product Growth Model: A Sensitivity Analysis for a High Technology Product". *Journal of Marketing*, (45), pp. 81-90.

Ullman, R. and Ferrera, D. (1998) "Crime on the Internet," *Boston Bar Journal*, Nov./Dec., no.6.

Wright, M. Upritchard, C. and Lewis, T. (1997) "A Validation of the Bass New Product Diffusion Model in New Zealand" *Marketing Bulletin*, (8), pp. 15-29.

ABOUT THE AUTHORS

Kallol K. Bagchi received his first Ph.D. in Computer Science from Jadavpur University, India, in 1988 and his second Ph.D. in Business from Florida Atlantic University in 2001. He taught at European universities for six years and also worked in industry for several years. His journal articles, conference proceedings articles, and books are in the computer science and MIS areas. He teaches MIS courses at The University of Texas at El Paso. His present research interests are in global information technology, adoption and diffusion of information technology, security, networking, and simulation. He is listed in the *International Who's Who of Professional Educators in 2002*. He is a member of the ACM, IEEECS, AIS, and the Decision Sciences Institute.

Godwin J. Udo is the Chairperson of the Department of Information and Decision Sciences and the Southwestern Bell Professor of Business in the College of Business Administration at The University of Texas at El Paso. He holds a Ph.D. in Industrial Management from Clemson University. He teaches Information Systems courses at graduate and undergraduate levels. He is the author of over 80 research articles in the areas of technology transfer, computer security, IT adoption and management, and decision support systems. He is a member of Decision Sciences Institute, AITP, APICS, AIS, and INFORMS.

Communications of the Association for Information Systems

ISSN: 1529-3181